

PCTWELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales BüroINTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

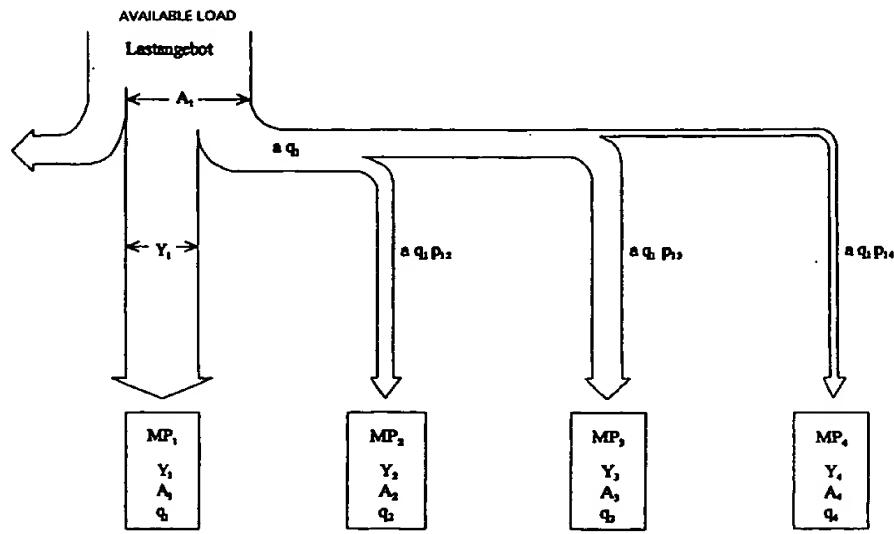
(51) Internationale Patentklassifikation ⁷ : G06F 9/46	A1	(11) Internationale Veröffentlichungsnummer: WO 00/43874 (43) Internationales Veröffentlichungsdatum: 27. Juli 2000 (27.07.00)
(21) Internationales Aktenzeichen: PCT/EP00/00317		(81) Bestimmungsstaaten: CN, JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(22) Internationales Anmeldedatum: 17. Januar 2000 (17.01.00)		
(30) Prioritätsdaten: 99101122.2 21. Januar 1999 (21.01.99) EP		Veröffentlicht <i>Mit internationalem Recherchenbericht. Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>
(71) Anmelder (<i>für alle Bestimmungsstaaten ausser US</i>): SIEMENS AKTIENGESELLSCHAFT [DE/DE]; Wittelsbacherplatz 2, D-80333 München (DE).		
(72) Erfinder; und		
(75) Erfinder/Anmelder (<i>nur für US</i>): ROSS, Christopher [GB/DE]; Kirchenstr. 89, D-81675 München (DE). HANSELKA, Peter [DE/DE]; Pidinger Str. 13, D-81379 München (DE). RAICHLE, Gabriele [DE/DE]; Ambacherstr. 39, D-81476 München (DE).		
(74) Gemeinsamer Vertreter: SIEMENS AKTIENGESELLSCHAFT; Postfach 22 16 34, D-80506 München (DE).		

(54) Title: LOAD DISTRIBUTION METHOD FOR A MULTIPROCESSOR SYSTEM AND CORRESPONDING MULTIPROCESSOR SYSTEM

(54) Bezeichnung: LASTVERTEILUNGSVERFAHREN EINES MULTIPROZESSORSYSTEMS UND MULTIPROZESSORSYSTEM

(57) Abstract

The invention relates to a method for distributing a load in a real time multiprocessor system and to a corresponding multiprocessor system. A distribute quota establishing the proportion of the distributable load to actually be distributed is determined on each processor. This distribute quota is redetermined at intervals. The only information which is required by the processors for each interval of time are load value indicators, which depend on an estimated load. Probabilities indicating how the load is transferred from one processor to the others during the load distribution are also determined. Each processor then distributes its distributable load when its distribute quota exceeds a predetermined value, in accordance with said distribute quota and its load distribution factors.

**Best Available Copy**

(57) Zusammenfassung

Die Erfindung betrifft ein Verfahren zur Lastverteilung in einem Realzeit-Multiprozessorsystem und ein Multiprozessorsystem, wobei auf jedem Prozessor eine Verteilquote geführt wird, die den Anteil der verteilbaren Last, der tatsächlich verteilt werden soll, festsetzt. Die Verteilquote wird in Zeitintervallen neu bestimmt. Die einzige Information, die jedes Zeitintervall von den anderen Prozessoren benötigt wird, sind Lastwertindikatoren, die von einer geschätzten Last abhängen. Zusätzlich werden Wahrscheinlichkeiten geführt, welche angeben, wie bei Lastverteilung Last von einem auf die anderen Prozessoren übertragen wird. Anschliessend verteilt jeder Prozessor anhand seiner Verteilquote und seiner Lastverteilungsfaktoren seine verteilbare Last an andere Prozessoren, wenn seine Verteilquote einen vorgegebenen Wert überschreitet.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		

Beschreibung

Lastverteilungsverfahren eines Multiprozessorsystems und
Multiprozessorsystem

5

Die Erfindung betrifft ein Verfahren zur Lastverteilung in einem Multiprozessorsystem, insbesondere in einem Multiprozessorsystem eines Kommunikationssystems, bei dem anfallende Aufgaben von mehreren Prozessoren MP_i (mit $i=1, 2, \dots, n$) unter Realzeitbedingungen abgearbeitet werden können und ein Multiprozessorsystem, insbesondere eines Kommunikationssystems, mit einem Lastverteilungsmechanismus.

Ein ähnliches Verfahren zur Lastverteilung in einem

15 Multiprozessorsystem, insbesondere in einem Multiprozessorsystem eines Kommunikationssystems, ist beispielsweise aus der Europäischen Patentanmeldung EP 0 645 702 A1 der Anmelderin bekannt. Diese Schrift offenbart ein Verfahren zum Lastenausgleich in einem Multiprozessorsystem, 20 insbesondere ein Multiprozessorsystem eines Kommunikationssystems, bei dem anfallende Aufgaben von mehreren Prozessoren unter Realzeitbedingungen abgearbeitet werden können, wobei zur Durchführung des Lastausgleiches allgemein die folgenden Verfahrensschritte genannt sind:

25 - jeder Prozessor ermittelt seinen Lastzustand in Form einer quantifizierten Größe,
- jedem Prozessor werden die Lastzustände der anderen Prozessoren innerhalb eines Zeitrasters mitgeteilt,
- jeder Prozessor gibt in Abhängigkeit vom Überschreiten 30 einer bestimmten Größe seines Lastzustandes und in Abhängigkeit von den Lastzuständen der übrigen Prozessoren zumindest einen Teil der bei ihm anfallenden Aufgaben an die übrigen Prozessoren ab, und

2

- die abgegebenen Aufgaben werden entsprechend den Lastzuständen der übrigen Prozessoren auf diese aufgeteilt.

5 Im Ausführungsbeispiel wird das Verfahren dahingehend konkretisiert, daß im Betrieb ständig und vor dem Einstieg in die Lastverteilung, die hier erst ab dem Erreichen einer bestimmten Überlast beginnt, Verteilungsquoten errechnet werden, nach denen die einzelnen Prozessoren im Fall der Überlast an andere Prozessoren ihre verteilbare Last abgeben.

10 10 Ist das System dauerhaft ungleichmäßig ausgelastet, so wird die Last erst bei Überlast eines oder mehrerer Prozessoren verteilt. Damit geht aber unnötige Lastabwehr einher. Die Reduktion der Überlastschwelle auf einen niedrigeren Wert führt zu keinem befriedigenden Ergebnis, weil dann unnötig

15 15 viel Last verteilt wird und es zu Schwingungszuständen kommen kann. Diese Situation ergibt sich aus der dort getroffenen Annahme, daß die Überlast oder die ungleichmäßige Belastung von kurzer Dauer ist.

Es ist daher Aufgabe der Erfindung, ein verbessertes

20 20 Lastverteilungsverfahren für ein Multiprozessorsystem anzugeben, welches rechtzeitig und "weich" einsetzt und dadurch dauerhafte Schieflastzustände im Lastangebot ohne Lastabwehr bewältigt. Außerdem soll auch ein entsprechendes Multiprozessorsystem angegeben werden.

25 Die Aufgabe wird einerseits durch ein Verfahren mit den Verfahrensschritten des ersten Verfahrensanspruches und andererseits durch ein Multiprozessorsystem mit den Merkmalen des ersten Vorrichtungsanspruches gelöst.

30 30 Demgemäß schlagen die Erfinder ein Verfahren zur Lastverteilung in einem Multiprozessorsystem, insbesondere in einem Multiprozessorsystem eines Kommunikationssystems, bei dem anfallende Aufgaben von mehreren Prozessoren MP_i (mit

35 35 $i=1, 2, \dots, n$) unter Realzeitbedingungen abgearbeitet werden

3

können, mit folgenden iterativen und sich in Zeitintervallen CI wiederholenden Verfahrensschritten vor:

- jeder Prozessor MP_i ermittelt seinen tatsächlichen Lastzustand Y_i - bestimmt gegebenenfalls direkt hieraus einen mehrwertigen Laststatus (load state) MP_{ls_i} - und schätzt in Abhängigkeit von zuvor mitgeteilten Verteilungsquoten $q_i(\text{alt})$ (mit $q_i = \text{an andere Prozessoren } MP_k \text{ nach Möglichkeit zu verteilender Lastanteil}$) und dem typischerweise verteilbaren Anteil V einer typischen Aufgabe seine angebotene Last A_i , die zu einem mehrwertigen Lastindikationswert (Balancing Indicator) MP_{bi_i} führt,
- jeder Prozessor MP_i teilt seinen Lastindikationswert MP_{bi_i} den jeweils anderen Prozessoren MP_k (mit $k=1, 2, \dots, i-1, i+1, \dots, n$) mittelbar oder unmittelbar mit,
- jeder Prozessor MP_i bestimmt seine Lastverteilungsfaktoren p_{ij} (mit $j=1, 2, \dots, n$) in Abhängigkeit von den Lastindikationswerten MP_{bi_k} dieser anderen Prozessoren MP_k ,
- jeder Prozessor MP_i bestimmt seine Verteilungsquote $q_i(\text{neu})$ in Abhängigkeit von seinem tatsächlichen Lastzustand Y_i und den Lastverteilungsfaktoren p_{ij} ,
- jeder Prozessor MP_i verteilt anhand seiner Quote q_i und seiner Lastverteilungsfaktoren p_{ij} seine verteilbare Last an andere Prozessoren MP_k , wenn seine Verteilungsquote $q_i(\text{neu})$ einen vorgegebenen Wert q_v überschreitet.

Zur Abschätzung der angebotenen Last A_i eines Prozessors MP_i ist es vorteilhaft, die Formel $A_i := Y_i / (1 - q_i V)$ zu verwenden.

Vorteilhaft ist auch eine Unterteilung des mehrwertigen Lastindikationswertes (balancing indicator) MP_{bi_i} in drei

4

diskrete Werte, wobei vorzugsweise die folgende Abgrenzung mit Schwellenwerten gilt: NORMAL für MP_{b*i*}, wenn die Prozessorauslastung 0 bis 70% beträgt, HIGH für MP_{b*i*}, wenn die Prozessorauslastung 70% bis 85% beträgt, und OVERLOAD für 5 MP_{b*i*}, wenn die Prozessorauslastung über 85% beträgt.

Vorteilhaft ist es auch, wenn eine Hysterese bei Lastzustandswechsel aufgrund von Schwellenwertüberschreitung oder Schwellenwertunterschreitung bei steigender oder 10 fallender Prozessorauslastung eingeführt wird.

Außerdem kann es vorteilhaft sein, wenn der Lastindikationswert (balancing indicator) MP_{b*i*} bezüglich Änderungen einer zeitlichen Hysterese unterliegt und damit 15 eine gewisse Trägheit erfährt. Als Hysteresegrenze können vorteilhaft Werte von 1 bis 2 Zeitintervallen CI angenommen werden.

Bezüglich des mehrwertigen Laststatus (load state) MP_{l*i*} wird 20 als besonders bevorzugt die Annahme von vier diskreten Werten vorgeschlagen, wobei vorzugsweise angenommen wird: NORMAL für MP_{l*i*}, wenn die Prozessorauslastung unter 70% liegt, HIGH für MP_{l*i*}, wenn die Prozessorauslastung 70% bis 85%, OVERLOAD für MP_{l*i*}, wenn die Prozessorauslastung über 85% liegt und 25 EXTREME für MP_{l*i*}, wenn dauerhaft der Lastzustand OVERLOAD vorherrscht. Auch hier kann es vorteilhaft sein, wenn der Laststatus (load state) MP_{l*i*} bezüglich Änderungen einer Hysterese unterliegt. Als Hysteresegrenze können vorteilhaft Werte von 1 bis 2 Zeitintervallen CI angenommen werden.

30

Weitere vorteilhafte Annahmen bei der Durchführung des erfindungsgemäßen Verfahrens sind: der typische verteilbare Anteil V einer typischen Aufgabe soll der durchschnittliche

5

oder maximale Anteil sein und als typische Bearbeitungszeit einer Aufgabe soll eine durchschnittliche oder maximale Bearbeitungszeit einer Aufgabe angenommen werden.

Vorteilhaft kann hierbei der jeweilige Durchschnittswert oder

- 5 Maximalwert eines Anteils beziehungsweise einer Aufgabe auch während der Betriebszeit ständig ermittelt und gegebenenfalls als gleitender Wert mitgeführt und aktualisiert in das Lastverteilungsverfahren übernommen werden. Günstig ist es hierbei, wenn die Zeitdauer, über die die gleitenden Werte
- 10 ermittelt werden, groß gegenüber dem Kontrollintervall CI ist.

Besonders vorteilhaft ist es auch, wenn für den vorgegebenen Wert q_v der Verteilungsquote q_i , ab dem der Prozessor MP_i

- 15 verteilbare Last an andere Prozessoren MP_k verteilt, gilt:
 $0,05 < q_v < 0,3$, vorzugsweise $0,1 < q_v < 0,25$, vorzugsweise $q_v = 0,2$.

Weiterhin kann das erfundungsgemäße Verfahren besonders vorteilhaft ausgestaltet werden, wenn bei der Berechnung der

- 20 Verteilungsquote q_i die folgenden Kriterien erfüllt werden:

- $p_{ii} := 0$
- falls MP_{bi,j} einer mittlere Last entspricht, vorzugsweise MP_{bi,j}=NORMAL, gilt: $p_{ij}(\text{neu}) = p_{ij}(\text{alt}) + p_{c1}/n$, für $j=1, \dots, n$ und $i \neq j$
- 25 - falls MP_{bi,j} einer hohen Last entspricht, vorzugsweise MP_{bi,j}=HIGH gilt: $p_{ij}(\text{neu}) = p_{ij}(\text{alt}) - p_{c2}/n$, für $j=1, \dots, n$ und $i \neq j$
- falls MP_{bi,j} einer Überlast entspricht, vorzugsweise MP_{bi,j}=OVERLOAD, gilt: $p_{ij}(\text{neu}) = 0$
- 30 - wobei vorzugsweise die p_{ij} ($j=1, \dots, n$) mit der Summe p_{sum} der p_{ij} auf 1 normiert wird und

6

- als Initialisierungswert beim Beginn der Verteilungsprozesse alle p_{ij} , ausgenommen p_{ii} , gleich sind.

5 Als vorteilhafte Zahlenwerte können für die Konstante p_{c1} $0,1 < p_{c1} < 0,5$, vorzugsweise $0,2 < p_{c1} < 0,3$ und vorzugsweise $p_{c1}=0,25$ angenommen werden. Ebenso ist es vorteilhaft für die Konstante p_{c2} $0,1 < p_{c2} < 0,5$, vorzugsweise $0,2 < p_{c2} < 0,3$, vorzugsweise $p_{c2}=0,25$ zu setzen. Auch kann der
10 Initialisierungswert der p_{ij} beim Beginn der Verteilungsprozesse gleich $(n-1)^{-1}$ gesetzt werden.

Weiterhin kann das erfindungsgemäße Verfahren besonders
15 vorteilhaft ausgestaltet werden, wenn bei der Berechnung der Lastindikationswerte $MPbi_i$ die folgenden Kriterien erfüllt werden:

- falls $MPbi_i$ der höchsten Last entspricht, vorzugsweise $MPbi_i=EXTREME$, gilt: $q_i(\text{neu})=q_{c1}$,
- 20 - falls $p_{sum} \geq 1$ gilt:
 - falls der tatsächliche Lastzustand Y_i größer als ein vorgegebener Wert $threshold_H$ ist, wird q_i vergrößert mit $q_i=\min\{q_i+c_{q1}, 1\}$,
 - falls der tatsächliche Lastzustand Y_i kleiner als ein vorgegebene Wert $threshold_N$ ist, wird q_i verkleinert mit
25 $q_i=\max\{q_i-c_{q2}, c_{q3}\}$, mit $0 < c_{q3} < q_v$, vorzugsweise $c_{q3}=0,1$,
 - andernfalls ($threshold_N \leq Y_i \leq threshold_H$) erhält q_i einen Zwischenwert zwischen den beiden oben genannten Alternativen, vorzugsweise durch lineare Interpolation
- 30 - falls $p_{sum} \leq 1$ gilt: $q_i(\text{neu})=q_i(\text{alt}) * p_{sum}$.

Zur optimalen Ausgestaltung des Verfahrens sind für die Konstante c_{q1} die folgenden Zahlenbereiche und -werte

7

bevorzugt: $0,05 < c_{q1} < 0,3$, vorzugsweise $0,1 < c_{q1} < 0,2$, vorzugsweise $c_{q1}=0,15$. Außerdem kann für die Konstante c_{q2} vorzugsweise $0,05 < c_{q2} < 0,2$, vorzugsweise $c_{q2}=0,10$ angenommen werden.

5

Bezüglich der Konstanten threshold_N gilt als bevorzugter Wertebereich: $0,6 < \text{threshold}_N < 0,8$, vorzugsweise $\text{threshold}_N = 0,7$.

10 Bezüglich der Konstanten threshold_H gilt als bevorzugter Wertebereich: $0,7 < \text{threshold}_H < 0,95$, vorzugsweise $\text{threshold}_H = 0,85$.

15 Eine andere Ausgestaltung des erfindungsgemäßen Verfahrens sieht vor, daß zusätzlich in jedem Zeitintervall CI ein Überlastwert OL_i der Prozessoren MP_i ermittelt wird, der ein Maß für die Größe der Überlast ist und als Maßstab zur Überlastabwehr dient, mit $OL_i=0,1, \dots, m$, und die Verteilungsquote q_i auf jeden Fall vergrößert wird, falls
20 $OL_i > 0$ mit $q_i(\text{neu}) := \min\{q_i(\text{alt}) + c_{q1}, 1\}$.

Erfindungsgemäß besteht auch die Möglichkeit, eine Adaption des Lastverteilungsverfahrens an sich ändernde Randbedingungen durchzuführen, indem die oben angegebenen
25 Konstanten ($q_v, p_{c1}, p_{c2}, q_{c1}, q_{c2}, \text{threshold}_N, \text{threshold}_H, c_{q1}, c_{q2}, c_{q3}$) im Betrieb zumindest teilweise angepaßt werden.

30 Erfindungsgemäß wird außerdem ein Multiprozessorsystem, insbesondere eines Kommunikationssystems, mit mehreren Prozessoren MP_i (mit $i=1, 2, \dots, n$) zur Ausführung anfallender Aufgaben unter Realzeitbedingungen, vorgeschlagen, wobei:
- jeder Prozessor MP_i Mittel aufweist, um seinen tatsächlichen Lastzustand Y_i zu bestimmen, -

gegebenenfalls direkt hieraus einen mehrwertigen Laststatus (load state) MP_{ls_i} zu bestimmen - und in Abhängigkeit von zuvor mitgeteilten Verteilungsquoten q_i (alt) (mit q_i =an andere Prozessoren MP_k nach Möglichkeit zu verteilender Lastanteil) und dem typischerweise verteilbaren Anteil V einer typischen Aufgabe seine angebotene Last A_i zu schätzen, die zu einem mehrwertigen Lastindikationswert (Balancing Indicator) MP_{bi_i} führt,

10 - jeder Prozessor MP_i Mittel aufweist, um seinen Lastindikationswert MP_{bi_i} den jeweils anderen Prozessoren MP_k (mit $k=1,2,\dots,i-1,i+1,\dots,n$) mittelbar oder unmittelbar mitzuteilen,

- jeder Prozessor MP_i Mittel aufweist, um seine Lastverteilungswahrscheinlichkeiten p_{ij} (mit $j=1,2,\dots,n$) in Abhängigkeit von den Lastindikationswerten MP_{bi_k} dieser anderen Prozessoren MP_k zu bestimmen,

- jeder Prozessor MP_i Mittel aufweist, um seine Verteilquote q_i (neu) in Abhängigkeit von seinem tatsächlichen Lastzustand Y_i zu bestimmen, und

- jeder Prozessor MP_i Mittel aufweist, um anhand seiner Quote q_i und seiner Lastverteilungsfaktoren p_{ij} seine verteilbare Last an andere Prozessoren MP_k zu verteilen, wenn seine Verteilquote q_i (neu) einen vorgegebenen Wert q_v überschreitet.

Erfnungsgemäß kann das oben vorgeschlagene Multiprozessorsystem so ausgestaltet werden, daß jeweils eines der oben genannten Verfahren implementiert ist, wobei die Implementierung durch eine entsprechende Programmierung der Prozessoren erfolgt.

9

Es ist noch darauf hinzuweisen, daß der Index (alt) sich jeweils auf die Werte des vorhergehenden Iterationsschrittes, beziehungsweise der Index (neu) sich auf den jetzt aktuellen Iterationsschritt beziehen.

5

Der besondere Vorteil des erfindungsgemäßen Verfahrens und des entsprechenden Multiprozessorsystems liegt darin, daß es im Gegensatz zum eingangs genannten Stand der Technik einen "weichen" Einstieg in die Lastverteilung gewährleistet und dadurch anpassungsfähiger, weniger anfällig gegen Schieflastsituationen ist und Schwingungszustände besser vermieden werden. Im Endeffekt wird hierdurch die Wahrscheinlichkeit für die Abwehr von Aufgaben, insbesondere Vermittlungsaufgaben reduziert.

10

Weitere Ausgestaltungen, zusätzliche Merkmale und Vorteile der Erfindung ergeben sich aus der nachfolgenden Beschreibung eines bevorzugten Ausführungsbeispiels unter Bezugnahme auf die Zeichnungen.

15

Es versteht sich, daß die vorstehend genannten und nachstehend noch zu erläuternden Merkmale der Erfindung nicht nur in der jeweils angegebenen Kombination, sondern auch in anderen Kombinationen oder in Alleinstellung verwendbar sind, ohne den Rahmen der Erfindung zu verlassen.

Die Figuren zeigen im einzelnen:

20

Figur 1: Flußbild des anfallenden und verteilten Lastangebotes

Figur 2a: Grafische Darstellung der Entscheidungen zur Aktualisierung der Lastverteilungsfaktoren p_{ij}

Figur 2b: Grafische Darstellung der Entscheidungen zur Aktualisierung der Verteilungsquoten q_i

25

Figur 3: Formel zur linearen Interpolation von q_i

10

Das erfindungsgemäße Verfahren (Normal Load Balancing=NLB) ist ein quotiertes Load Balancing Verfahren, das auf einem Multiprozessorsystem, insbesondere in einer Vermittlungsstelle eines Kommunikationssystems, zur Verteilung anfallender Arbeitslasten auf die jeweils anderen Prozessoren abläuft und sicherstellen soll, daß andauernde Schieflastsituationen bewältigt und möglichst alle angeforderten Aufgaben in möglichst kurzer Zeit abgearbeitet werden. Nachfolgend soll eine besonders vorteilhafte Ausführungsform dieses Verfahrens beschrieben werden.

Auf jedem Prozessor MP_i mit $i=1, 2, \dots, n$ wird eine Verteilquote q_i geführt, die den Anteil V der verteilbaren Last, der tatsächlich verteilt werden soll, festsetzt. Eine solche Quote ermöglicht einen weicheren Ein- beziehungsweise Ausstieg aus der Lastverteilung an andere Prozessoren. Auf diese Weise werden Schwingzustände und Lastschwankungen vermieden. Dies kann beispielsweise der Fall sein, wenn ein Prozessor so viel Last an einen anderen verteilt, daß dieser wiederum überlastet wird.

Die Verteilquote q_i wird jedes Zeitintervall CI neu bestimmt. Die einzige Information, die jedes CI von den anderen Prozessoren MP_k mit $k=1, \dots, i-1, i+1, \dots, n$ benötigt wird, sind Lastwertindikatoren (Balancing Indikatoren) $MPbi_i$. Diese Lastwertindikatoren sind - ähnlich wie die Laststatuswerte (Load States) von der Load Control - Lastzustände mit den Wertigkeiten NORMAL, HIGH oder OVERLOAD. Während der Load State anhand der tatsächlich bearbeiteten Last Y_i des Prozessors MP_i bestimmt wird, wird der Lastwertindikatoren $MPbi_i$ aus einer Schätzung der aktuell angebotenen Last A_i ermittelt. Die geschätzte angebotene Last A_i kann durch Lastverteilung erheblich mehr als die tatsächlich bearbeitete Last Y_i sein und stellt die maßgebliche Größe dar, die (in Form des Lastwertindikators $MPbi_i$) ein Prozessor MP_i den anderen MP_k als Information zur Verfügung stellt.

Zusätzlich zur Verteilungsquote q_i werden auf jedem MP_i Wahrscheinlichkeiten p_{ij} , geführt, welche die Wahrscheinlichkeit angeben, daß bei Lastverteilung Last vom i-ten Prozessor MP_i auf den j-ten Prozessor MP_j übertragen wird. Die Wahrscheinlichkeiten werden so bestimmt, daß, wenn etwa der j-te Prozessor MP_j schon viel Last zu bearbeiten hat und deshalb nur wenig zusätzliche Last aufnehmen kann, das zugehörige p_{ij} kleiner ist als das P_{ik} für einen freien MP_k .

In Figur 1 wird das Zusammenspiel der p_{ij} und q_i veranschaulicht. Die doppelte Indizierung "ij" der Kenngrößen besagt, daß jeweils der Prozessor mit der Nummer des ersten Index (hier i) jeweils eine "Spalte" von n Werten mit dem zweiten Index (hier j) kennt. Es ist zu bemerken, daß jeder Prozessor nur seine relevanten Werte (also seine Spalte) kennt, wobei insgesamt im System eine quadratische Matrix bekannt ist. So ist zum Beispiel p_{ij} die Wahrscheinlichkeit, daß Last vom i-ten MP auf den j-ten MP verteilt wird, wenn der i-te MP zu viel Last hat.

In der Figur 1 ist außerdem die tatsächlich bearbeitete Last des j-ten Prozessors MP_j mit Y_j , die geschätzte angebotene Last mit A_j und der Teil des Lastangebotes, der verlagert werden kann, mit a bezeichnet. Die gezeigte Lastsituation ist Überlast (OVERLOAD) auf MP_1 , auf den MP_k mit $k=2,3,4$ ist noch Raum für zusätzliche Aufgaben. Es wird gezeigt, wie der MP_1 einen ersten Teil der Last selbst bearbeitet und den Rest a verteilt. Von diesem Rest a geht der größte Anteil an MP_3 , der kleinste Anteil an MP_4 , der in diesem Beispiel also schon viel eigene Last zu bearbeiten hat. Nicht eingezzeichnet sind die Lasten, welche die MP_k außer von MP_1 noch erhalten. Die Breite der Fließbalken stellen ein Maß für die Größe der Last dar.

Gemäß dem Erfindungsgedanken ergibt sich also der folgende Algorithmus: Meldet der j-te Prozessor MP_j den Balancing Indikator NORMAL, wird auf dem jeweils betrachteten MP_i das

12

p_{ij} vergrößert. Es steigt also die Wahrscheinlichkeit, daß dieser Prozessor MP_i Last an MP_j abgibt, wenn er Last verteilen muß. Wird der Balancing Indikator HIGH gemeldet, so wird das p_{ij} verkleinert. Wird der Balancing Indikator 5 OVERLOAD gemeldet, wird p_{ij} auf Null gesetzt, so daß keine Last an den j-ten Prozessor MP_j abgegeben wird. Die Verteilungsquote q_i wird anschließend an die Bestimmung der p_{ij} verändert. Konnten viele der p_{ij} vergrößert werden, so ist die Summe der p_{ij} über j größer 1 und offenbar noch Platz auf 10 den anderen Prozessoren MP_k . Die Verteilungsquote q_i kann also nach den Erfordernissen des (betrachteten) Prozessors verändert werden.

Bei hoher Last Y_i auf dem betrachteten Prozessor MP_i wird die Verteilungsquote q_i vergrößert, bei niedriger Last wird q_i 15 verkleinert. Sind viele der p_{ij} verringert worden, dann ist die Summe der p_{ij} über j kleiner 1 und die Verteilungsquote q_i muß verringert werden.

Eine Veranschaulichung dieser Entscheidungen ist in den 20 Figuren 2a und 2b dargestellt. Die Entscheidungsdiagramme zeigen die Aktualisierungsalgorithmen für p_{ij} (Figur 2a) und für Verteilungsquote q_i (Figur 2b), die jedes Zeitintervall CI für den i-ten Prozessor MP_i durchgeführt werden.

25 Bei dem erfindungsgemäßen Lastverteilungsverfahren (NLB) werden einige Parameter (Konstanten) benötigt, deren Wahl das Verhalten in bestimmten Lastsituationen stark beeinflussen kann. Es ergibt sich in den meisten Fällen ein Konflikt zwischen einem Lastverteilungsverfahren, das schnell auf Laständerungen reagieren kann, und einem stabilen 30 Lastverteilungsverfahren, das nicht zu Schwingungen und zum Weiterverteilen von Aufgaben neigt. "Weiterverteilen" bedeutet hier das gleichzeitige Verteilen von eigener Last und das Bearbeiten von fremder Last auf einem Prozessor.

35 Folgende Parameterveränderungen bewirken ein schneller reagierendes NLB:

13

- Das stärkere Verändern von q_i mit: $0,15 < c_{q1}, 0,1 < c_{q2}$
- Das stärkere Verändern der p_{ij} mit: $0,25 < p_{c1}, 0,25 < p_{c2}$
- Das spätere Setzen der Lastindikationswerte $MPbi_i$ mit:
 $threshold_H > 0,7$ (d.h. erst bei höherer Last 'HIGH' an die anderen Prozessoren MP_k melden)

5

Detailliert verläuft das bevorzugte Verfahren bei einem Multiprozessor-Kommunikationsrechner also wie folgt:

10 Als Dauer des Zeitintervalls (Kontrollintervall) CI des Zeitrasters, mit dem das Verfahren iterativ abläuft, wird bei den derzeit bekannten Multiprozessorsystemen der Vermittlungstechnik bevorzugt 1 bis 2 Sekunden gewählt. Es ist selbstverständlich, daß mit steigender Prozessorleistung
15 das Zeitintervall gekürzt werden kann.

Jedes Kontrollintervall CI werden die Größen q_i , p_{ij} , MP_{ls_i} und MP_{bi_j} aktualisiert.

20 Die tatsächlich bearbeitete Last Y_i eines Prozessors MP_i wird als Prozessorlaufzeitgröße, gemessen in Erlang, ermittelt.

25 Die geschätzte angebotene Last A_j eines Prozessors MP_i wird aus der Verteilquote q_i des aktuellen Kontrollintervalls CI und dem geschätzten verteilbaren Anteil einer durchschnittlichen Aufgabe, zum Beispiel der Abarbeitung eines Calls, ermittelt.

Es gilt:

30 Die Anzahl der Prozessoren MP_i im Multiprozessorsystem ist n .

$A_i := Y_i / (1 - q_i V)$, wobei V der verteilbare Anteil eines Calls ist.

35 MP_{ls_i} : Load State des i-ten MP, kann die Werte NORMAL, HIGH, OVERLOAD oder EXTREME annehmen. Zur Berechnung des Load States wird die tatsächlich bearbeitete Last Y_i herangezogen.

Zur Vermeidung von vorschnellen Änderungen des MP_{ls_i} werden Hysteresen eingeführt. Wird etwa der MP_{ls_i} von NORMAL auf HIGH gesetzt, muß $Y_i > threshold_N + \Delta_+$ sein, wohingegen, um von HIGH nach NORMAL zu kommen, $Y_i < threshold_N - \Delta_-$ sein

5 muß. Diese Vorgehensweise ist auch als High Water-Low Water Methode bekannt. Bei EXTREME muß aus systemtechnischen Gründen der Vermittlungsstelle das Verteilungsverfahren (Load Balancing) für diesen Prozessor MP_i abgeschaltet werden.

10 threshold_N: Ist die Normallastschwelle - nach Berücksichtigung einer Hysterese wird unterhalb dieser der MP_{ls} als NORMAL geführt, oberhalb als HIGH.

15 threshold_H: Hochlastschwelle - nach Berücksichtigung einer Hysterese und einer lastabhängigen zeitlichen Verzögerung (Startindikator) wird unterhalb dieser Schwelle der MP_{ls} als HIGH geführt, oberhalb als OVERLOAD.

Der Lastindikationswert (Balancing Indikator) MP_{b_i} des i-ten 20 Prozessors MP_i kann die Werte NORMAL, HIGH oder OVERLOAD annehmen. Dieser wird wie der MP_{ls_i} berechnet, nur wird hier anstelle der tatsächlichen Last Y_i die geschätzte angebotene Last A_i zugrundegelegt und andere Werte für Δ_+ und Δ_- genommen, mit $\Delta_+ = \Delta_- = 0.02$.

25 Zusätzlich wird ein Overload Level OL_i des Prozessors MP_i bestimmt, der die Werte 0... 6 annehmen kann und als Quantifizierung des Überlastzustandes des Prozessors MP_i gedacht ist. Ist der $OL_i > 0$, werden Calls abgewehrt, je höher 30 der Wert, desto wahrscheinlicher wird ein Call abgewiesen.

Die Last, die von MP_i nach MP_j verteilt werden soll, wird als Wahrscheinlichkeit p_{ij} ausgedrückt und kann somit Werte zwischen 0 und 1 annehmen.

35 Die Größe des Wertes p_{ij} bestimmt sich durch folgende Kriterien:

15

- Initialisiere p_{ij} mit $p_{ij} := (n-1)^{-1}$
- $p_{ii} := 0$, MP_i soll nicht an sich selbst verteilen.
- Falls MP_{bi,j}=NORMAL: $p_{ij} \rightarrow p_{ij} + 0.25/n$, $j=1, \dots, n$, $i \neq j$. Das alte p_{ij} kann vergrößert werden, weil auf dem Prozessor MP_j noch Platz ist.
- 5 - Falls MP_{bi,j}=HIGH: $p_{ij} \rightarrow p_{ij} - 0.25/n$. Das alte p_{ij} muß verkleinert werden, weil MP_j voll ausgelastet ist.
- Falls MP_{bi,j}=OVERLOAD: $p_{ij} = 0$. Es soll keine Last an überlastete Prozessoren MP_n abgeben werden.

10

Die neu bestimmten p_{ij} müssen noch normiert werden:

Setze $p_{sum} = \text{summe}(p_{ij})$ über $j=1, \dots, n$
und normiere (falls $p_{sum} > 0$) mit $p_{ij} \rightarrow p_{ij}/p_{sum}$

15

Anschließend wird die Verteilungsquote q_i mit den folgenden Kriterien bestimmt:

- Initialisierungswert: $q_i = 0.1$
- Falls der MP_{l,i} =EXTREME: $q_i = 0.1$. Dieser MP ist so stark überlastet, daß ihn auch der Eigenanteil für einen verteilten Call überfordern würde. Deshalb kein Load Balancing, sondern nur Abwehr; das Load Balancing ist zudem aus systemtechnischen Gründen der Vermittlungsstelle nicht sinnvoll.
- 20 - Falls $p_{sum} > 1$, kann offenbar mehr Last verteilt werden. Dann kann q_i nach den Erfordernissen des MP_i bestimmt werden, mit:
 1. Falls der $OL_i > 0$, q_i auf jeden Fall vergrößern mit: $q_i \rightarrow \min\{q_i + 0.15, 1\}$
 2. Falls $Y_i > threshold_H$, q_i vergrößern mit: $q_i \rightarrow \min\{q_i + 0.15, 1\}$
 - 30 3. Falls $Y_i < threshold_N$, q_i verkleinern mit: $q_i \rightarrow \max\{q_i - 0.10, 0.1\}$.
 4. Andernfalls, falls $threshold_N < Y_i < threshold_H$ gilt:

$$q_i \rightarrow \min\{\max\{q_i + (0.25 / (threshold_H - threshold_N)) * (Y_i - threshold_N) - 0.1, 0.1\}, 1.0\}$$
- 25 - Falls $p_{sum} < 1$, kann offenbar weniger Last verteilt werden. Dann kann q_i nach den Erfordernissen des MP_i bestimmt werden, mit:
 1. Falls der $OL_i < 0$, q_i auf jeden Fall verkleinern mit: $q_i \rightarrow \max\{q_i - 0.15, 0.1\}$
 2. Falls $Y_i < threshold_N$, q_i verkleinern mit: $q_i \rightarrow \max\{q_i - 0.15, 0.1\}$
 3. Falls $Y_i > threshold_H$, q_i vergrößern mit: $q_i \rightarrow \min\{q_i + 0.15, 1\}$.
 - 35 4. Andernfalls, falls $threshold_N < Y_i < threshold_H$ gilt:

$$q_i \rightarrow \min\{\max\{q_i + (0.25 / (threshold_H - threshold_N)) * (Y_i - threshold_N) + 0.1, 0.1\}, 1.0\}$$

16

Dies ist die lineare Interpolation zwischen der obigen Vergrößerung um 0.15 und der obigen Verkleinerung um 0.1. Die Formel ist nochmals in besser lesbbarer Weise in der Figur 3 dargestellt.

5

- Falls $p_{sum} < 1$, wurde offenbar zuviel Last verteilt und q_i muß verkleinert werden mit: $q_i \rightarrow q_i * p_{sum}$.
- Der Prozessor MP_i verteilt Last an andere Prozessoren MP_k , wenn $q_i > 0.25$ wird.

10

Das erfindungsgemäße Verfahren weist somit die folgenden Eigenschaften und Vorteile auf:

15

Ein sehr geringer Informationsoverhead zwischen den am Lastverteilungsverfahren beteiligten Prozessoren. Gegenseitig bekannt sind nur wenige, vorzugsweise dreiwertige Lastzustände, die nur einmal pro Kontrollintervall aktualisiert und verteilt werden.

20

Für jeden Prozessor gibt es eine Quote, die jedes Kontrollintervall aktualisiert wird und die den Anteil der Last regelt, die vom betrachteten Prozessor an die anderen beteiligten Prozessoren verteilt werden soll.

25

Für jeden Prozessor gibt es Einzelregulatoren, welche die zu verteilende Last auf die anderen Prozessoren aufteilen.

30

Das Verfahren ist nicht nur als "Feuerwehrmaßnahme" konzipiert, das erst wirksam wird, wenn ein Prozessor in Überlast gerät und gegebenenfalls Aufgaben (Calls) abgewehrt werden, sondern es setzt die Lastverteilung früher und weicher ein. Dadurch können Dauerschieflastzustände besser und mit weniger abgewiesenen Aufgaben (Calls) verarbeitet werden.

Im erfindungsgemäßen Verfahren werden die Lastzustände, die an die anderen Prozessoren verteilt werden, konsequent anhand der geschätzten angebotenen Last und nicht anhand der tatsächlich bearbeiteten Last ermittelt.

5

Das Verfahren benötigt kein Load Balancing Flag, das den Einstieg in die Lastverteilung regelt. Der Einstieg wird über die Verteilquote q_i geregelt. Weiterhin sind durch das Fehlen eines Load Balancing Flags gegenseitige Abhängigkeiten zwischen den Lastzuständen und dem Load Balancing Flag eliminiert worden. Dadurch ist eine nachträgliche Anpassung des Algorithmus an veränderte Bedingungen leichter möglich.

10

Die lastabhängige Veränderung der Einzelregulatoren (Lastverteilungsfaktoren p_{ij}) geschieht in Abhängigkeit von der Anzahl n der an der Lastverteilung beteiligten Prozessoren. Somit ist das Verfahren unabhängig von der Anzahl der beteiligten Prozessoren.

20

Die lastabhängige Veränderung der Verteilungsquoten und der Einzelregulatoren pro Kontrollintervall geschieht so, daß ein zu langsames "Heranschleichen" an den optimalen Wert vermieden wird.

25

Die lastabhängige Veränderung der Einzelregulatoren vermeidet ein Verharren der Werte auf die Einstellung der vorangegangenen Lastverteilungsperiode während einer Periode ohne Lastverteilung. Es wird vielmehr auf eine Ausgangsstellung zurückgeregelt.

30

Die aus dem Stand der Technik bekannte Trägheit in der Veränderung der Quoten wurde entfernt, um ein leichteres Nachführen an die tatsächlich vorliegende Lastsituation zu ermöglichen.

Patentansprüche

1. Verfahren zur Lastverteilung in einem
Multiprozessorsystem, insbesondere in einem
5 Multiprozessorsystem eines Kommunikationssystems, bei
dem anfallende Aufgaben von mehreren Prozessoren MP_i
(mit $i=1,2,\dots,n$) unter Realzeitbedingungen
abgearbeitet werden können, mit folgenden iterativen und
sich in Zeitintervallen CI wiederholenden
10 Verfahrensschritten:

- jeder Prozessor MP_i ermittelt seinen tatsächlichen
Lastzustand Y_i - bestimmt gegebenenfalls direkt
hieraus einen mehrwertigen Laststatus (load state)
 $MP_{ls,i}$ - und schätzt in Abhängigkeit von zuvor
15 mitgeteilten Verteilungsquoten $q_i(\text{alt})$ (mit $q_i=\text{an}$
andere Prozessoren MP_k nach Möglichkeit zu
verteilender Lastanteil) und dem typischerweise
verteilbaren Anteil V einer typischen Aufgabe
seine angebotene Last A_i , die zu einem mehrwertigen
20 Lastindikationswert (Balancing Indicator) $MP_{bi,i}$
führt,
- jeder Prozessor MP_i teilt seinen
Lastindikationswert $MP_{bi,i}$ den jeweils anderen
Prozessoren MP_k (mit $k=1,2,\dots,i-1,i+1,\dots,n$)
25 mittelbar oder unmittelbar mit,
- jeder Prozessor MP_i bestimmt seine
Lastverteilungswahrscheinlichkeiten p_{ij} (mit
 $j=1,2,\dots,n$) in Abhängigkeit von den
Lastindikationswerten $MP_{bi,k}$ dieser anderen
30 Prozessoren MP_k ,
- jeder Prozessor MP_i bestimmt seine Verteilungsquote
 $q_i(\text{neu})$ in Abhängigkeit von seinem tatsächlichen
Lastzustand Y_i und den Lastverteilungsfaktoren p_{ij}

19

- jeder Prozessor MP_i verteilt anhand seiner Quote q_i und seiner Lastverteilungsfaktoren p_{ij} seine verteilbare Last an andere Prozessoren MP_k wenn seine Verteilquote $q_i(\text{neu})$ einen vorgegebenen Wert q_v überschreitet.

5

2. Verfahren gemäß dem voranstehenden Anspruch 1, dadurch gekennzeichnet, daß die geschätzte angebotene Last A_i eines Prozessors MP_i nach der Formel $A_i := Y_i / (1 - q_i V)$ errechnet wird.

10

3. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß der mehrwertige Lastindikationswert (balancing indicator) $MPbi_i$ drei diskrete Werte, vorzugsweise NORMAL (=0 bis 0,7), HIGH (=0,7 bis 0,85) und OVERLOAD (=0,85 bis 1) annehmen kann.

15

4. Verfahren gemäß dem voranstehenden Anspruch 3, dadurch gekennzeichnet, daß der Lastindikationswert (balancing indicator) $MPbi_i$ bezüglich Änderungen einer Hysterese unterliegt.

20

5. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß der mehrwertige Laststatus (load state) $MPls_i$ vier diskrete Werte, vorzugsweise NORMAL (=0 bis 0,7), HIGH (=0,7 bis 0,85), OVERLOAD (=0,85 bis 1) und EXTREME (wenn Laststatus über mehrere CI OVERLOAD) annehmen kann.

25

6. Verfahren gemäß dem voranstehenden Anspruch 3, dadurch gekennzeichnet, daß der Laststatus (load state) $MPls_i$ bezüglich Änderungen einer Hysterese unterliegt.

30

7. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß der Wert der tatsächlichen Last Y_i proportional zur Prozessorlaufzeit ist.
5
8. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß als typischer verteilbarer Anteil V einer typischen Aufgabe CallP der durchschnittliche oder maximale Anteil ist.
10
9. Verfahren gemäß dem voranstehenden Anspruch 8, dadurch gekennzeichnet, daß der durchschnittliche oder maximale Anteil einer typischen Aufgabe ständig als gleitender Durchschnitt oder gleitender Maximalwert über eine vorgegebene Zeitspanne t_D ermittelt wird.
15
10. Verfahren gemäß dem voranstehenden Anspruch 9, dadurch gekennzeichnet, daß für die vorgegebene Zeitspanne t_D gilt: $t_D \gg CI$.
20
11. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß als typische Aufgabe eine durchschnittliche oder maximale Aufgabe angenommen wird.
25
12. Verfahren gemäß dem voranstehenden Anspruch 8, dadurch gekennzeichnet, daß die durchschnittliche oder maximale Aufgabe ständig als gleitender Durchschnitt oder gleitender Maximalwert über eine vorgegebene Zeitspanne t_D ermittelt wird.
30

21

13. Verfahren gemäß dem voranstehenden Anspruch 12, dadurch gekennzeichnet, daß für die vorgegebene Zeitspanne t_D gilt: $t_D \gg C_1$.

5 14. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß für den vorgegebenen Wert q_v der Verteilquote q_i , ab dem der Prozessor MP*i* verteilbare Last an andere Prozessoren MP_k verteilt gilt: $0,05 < q_v < 0,3$, vorzugsweise $0,1 < q_v < 0,25$, vorzugsweise $q_v = 0,2$.

10 15. Verfahren gemäß einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß die Berechnung der Verteilquote q_i die folgenden Kriterien erfüllt:

15 - $p_{ii} := 0$

- falls MP_{bij} einer mittlere Last entspricht, vorzugsweise MP_{bij}=NORMAL, gilt:
 $p_{ij}(\text{neu}) = p_{ij}(\text{alt}) + p_{c1}/n$, für $j=1, \dots, n$ und $i \neq j$

- falls MP_{bij} einer hohen Last entspricht, vorzugsweise MP_{bij}=HIGH gilt:
 $p_{ij}(\text{neu}) = p_{ij}(\text{alt}) - p_{c2}/n$, für $j=1, \dots, n$ und $i \neq j$

- falls MP_{bij} einer Überlast entspricht, vorzugsweise MP_{bij}=OVERLOAD, gilt:
 $p_{ij}(\text{neu}) = 0$

20 - wobei vorzugsweise die p_{ij} ($j=1, \dots, n$) mit der Summe p_{sum} der p_{ij} auf 1 normiert wird und

- als Initialisierungswert beim Beginn der Verteilungsprozesse alle p_{ij} , ausgenommen p_{ii} , gleich sind.

25 30 16. Verfahren gemäß dem voranstehenden Anspruch 15, dadurch gekennzeichnet, daß für die Konstante p_{c1} gilt:

22

$0,1 < p_{c1} < 0,5$, vorzugsweise $0,2 < p_{c1} < 0,3$, vorzugsweise $p_{c1}=0,25$.

17. Verfahren gemäß einem der voranstehenden Ansprüche 15-
5 16, dadurch gekennzeichnet, daß für die Konstante p_{c2} gilt: $0,1 < p_{c2} < 0,5$, vorzugsweise $0,2 < p_{c2} < 0,3$, vorzugsweise $p_{c2}=0,25$.

18. Verfahren gemäß einem der voranstehenden Ansprüche 15-
10 17, dadurch gekennzeichnet, daß der Initialisierungswert der p_{ij} beim Beginn der Verteilungsprozesse gleich $(n-1)^{-1}$ gesetzt wird.

19. Verfahren gemäß einem der voranstehenden Ansprüche 15-
15 18, dadurch gekennzeichnet, daß die Berechnung der Lastindikationswerte $MPbi_i$ die folgenden Kriterien erfüllt:

- falls $MPbi_i$ der höchsten Last entspricht, vorzugsweise $MPbi_i=EXTREME$, gilt:
20 $q_i(\text{neu})=q_{c1}$,
- falls $p_{sum} \geq 1$ gilt:
 - falls der tatsächliche Lastzustand Y_i größer als ein vorgegebener Wert $threshold_H$ ist, wird q_i vergrößert mit $q_i=\min\{q_i+c_{q1}, 1\}$,
 - falls der tatsächliche Lastzustand Y_i kleiner als ein vorgegebener Wert $threshold_N$ ist, wird q_i verkleinert mit $q_i=\max\{q_i-c_{q2}, c_{q3}\}$, mit $0 < c_{q3} < q_v$, vorzugsweise $c_{q3}=0,1$,
 - andernfalls ($threshold_N \leq Y_i \leq threshold_H$) erhält q_i einen Zwischenwert zwischen den beiden oben genannten Alternativen, vorzugsweise durch lineare Interpolation
- falls $p_{sum} \leq 1$ gilt: $q_i(\text{neu})=q_i(\text{alt}) * p_{sum}$.

20. Verfahren gemäß dem voranstehenden Anspruch 19, dadurch gekennzeichnet, daß für die Konstante c_{q1} gilt:
 $0,05 < c_{q1} < 0,3$, vorzugsweise $0,1 < c_{q1} < 0,2$, vorzugsweise
5 $c_{q1}=0,15$.

21. Verfahren gemäß einem der voranstehenden Ansprüche 19-
10 20, dadurch gekennzeichnet, daß für die Konstante c_{q2} gilt: $0,05 < c_{q2} < 0,2$, vorzugsweise $c_{q2}=0,10$.

22. Verfahren gemäß einem der voranstehenden Ansprüche 19-
15 21, dadurch gekennzeichnet, daß für die Konstante $threshold_N$ gilt: $0,6 < threshold_N < 0,8$, vorzugsweise
 $threshold_N = 0,7$.

15 23. Verfahren gemäß einem der voranstehenden Ansprüche 19-
20 22, dadurch gekennzeichnet, daß für die Konstante $threshold_H$ gilt: $0,7 < threshold_H < 0,95$, vorzugsweise
 $threshold_H = 0,85$

20 24. Verfahren gemäß einem der voranstehenden Ansprüche,
dadurch gekennzeichnet, daß zusätzlich ein
Überlastwert OL_i der Prozessoren MP_i ermittelt wird, der
ein Maß für die Größe der Überlast ist, mit $OL_i=0,1, \dots, m$
25 und die Verteilquote q_i auf jeden Fall vergrößert wird,
falls $OL_i > 0$ mit $q_i(\text{neu}) := \min\{q_i(\text{alt}) + c_{q1}, 1\}$.

25 25. Multiprozessorsystem, insbesondere eines
Kommunikationssystems, mit mehreren Prozessoren MP_i (mit
30 $i=1, 2, \dots, n$) zur Ausführung anfallender Aufgaben unter
Realzeitbedingungen, wobei:
- jeder Prozessor MP_i Mittel aufweist, um seinen
tatsächlichen Lastzustand Y_i zu bestimmen, -

24

gegebenenfalls direkt hieraus einen mehrwertigen
Laststatus (load state) MP_{ls_i} zu bestimmen - und in
Abhängigkeit von zuvor mitgeteilten
Verteilungsquoten q_i (alt) (mit $q_i = \text{an andere}$
5 Prozessoren MP_k nach Möglichkeit zu verteilender
Lastanteil) und dem typischerweise verteilbaren
Anteil V einer typischen Aufgabe, seine angebotene
Last A_i zu schätzen, die zu einem mehrwertigen
Lastindikationswert (Balancing Indicator) MP_{bi_i}
10 führt,

- jeder Prozessor MP_i Mittel aufweist, um seinen
Lastindikationswert MP_{bi_i} den jeweils anderen
Prozessoren MP_k (mit $k=1, 2, \dots, i-1, i+1, \dots, n$)
mittelbar oder unmittelbar mitzuteilen,
- jeder Prozessor MP_i Mittel aufweist, um seine
Lastverteilungswahrscheinlichkeiten p_{ij} (mit
 $j=1, 2, \dots, n$) in Abhängigkeit von den
Lastindikationswerten MP_{bi_k} dieser anderen
Prozessoren MP_k zu bestimmen,
- jeder Prozessor MP_i Mittel aufweist, um seine
Verteilquote q_i (neu) in Abhängigkeit von seinem
tatsächlichen Lastzustand Y_i zu bestimmen, und
- jeder Prozessor MP_i Mittel aufweist, um anhand
seiner Quote q_i und seiner Lastverteilungsfaktoren
 p_{ij} seine verteilbare Last an andere Prozessoren MP_k
zu verteilen, wenn seine Verteilquote q_i (neu) einen
25 vorgegebenen Wert q_v überschreitet.

30 26. Multiprozessorsystem gemäß Anspruch 25, dadurch
gekennzeichnet, daß eines der Verfahren gemäß einem der
Ansprüche 1-24 implementiert ist.

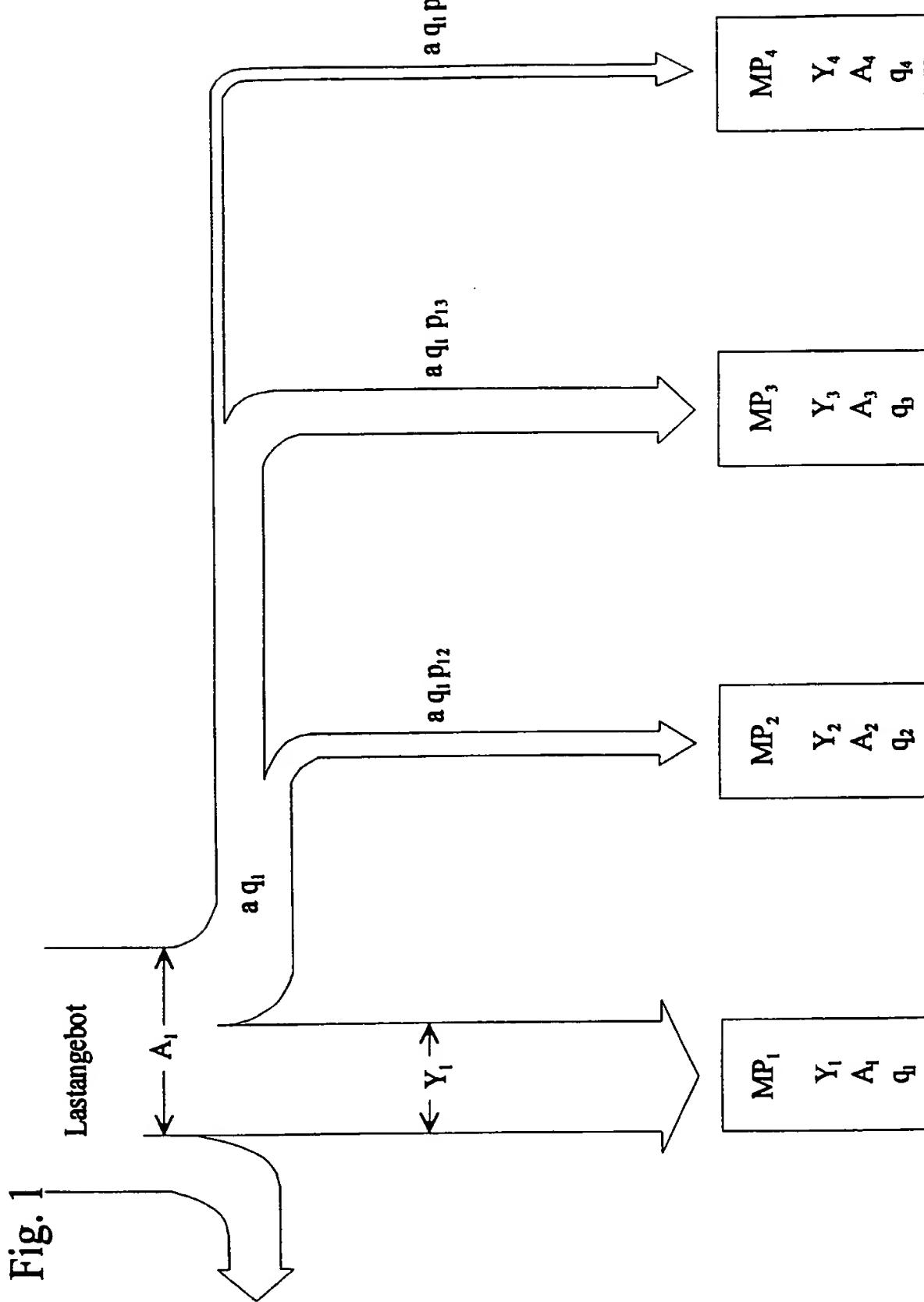
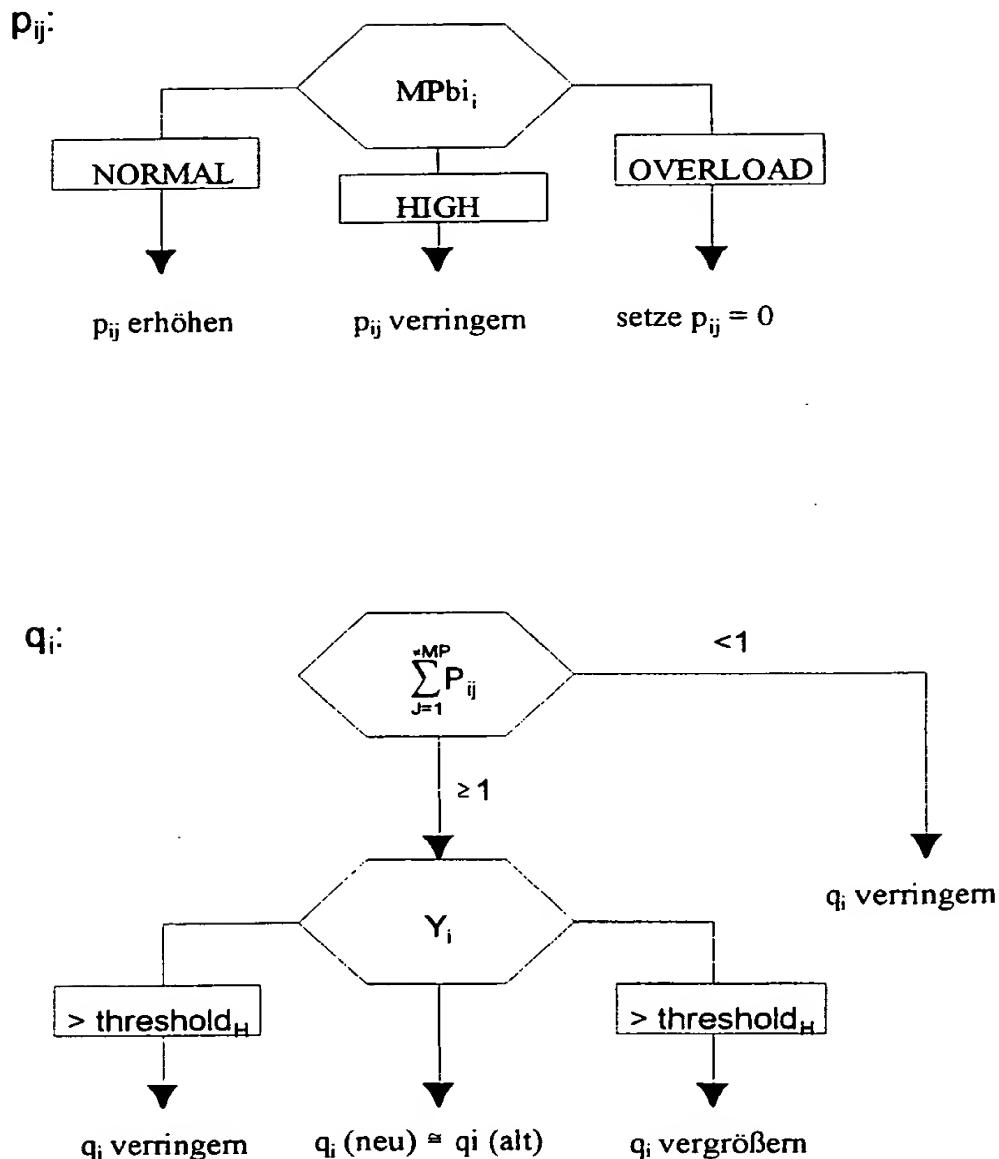




Fig. 2



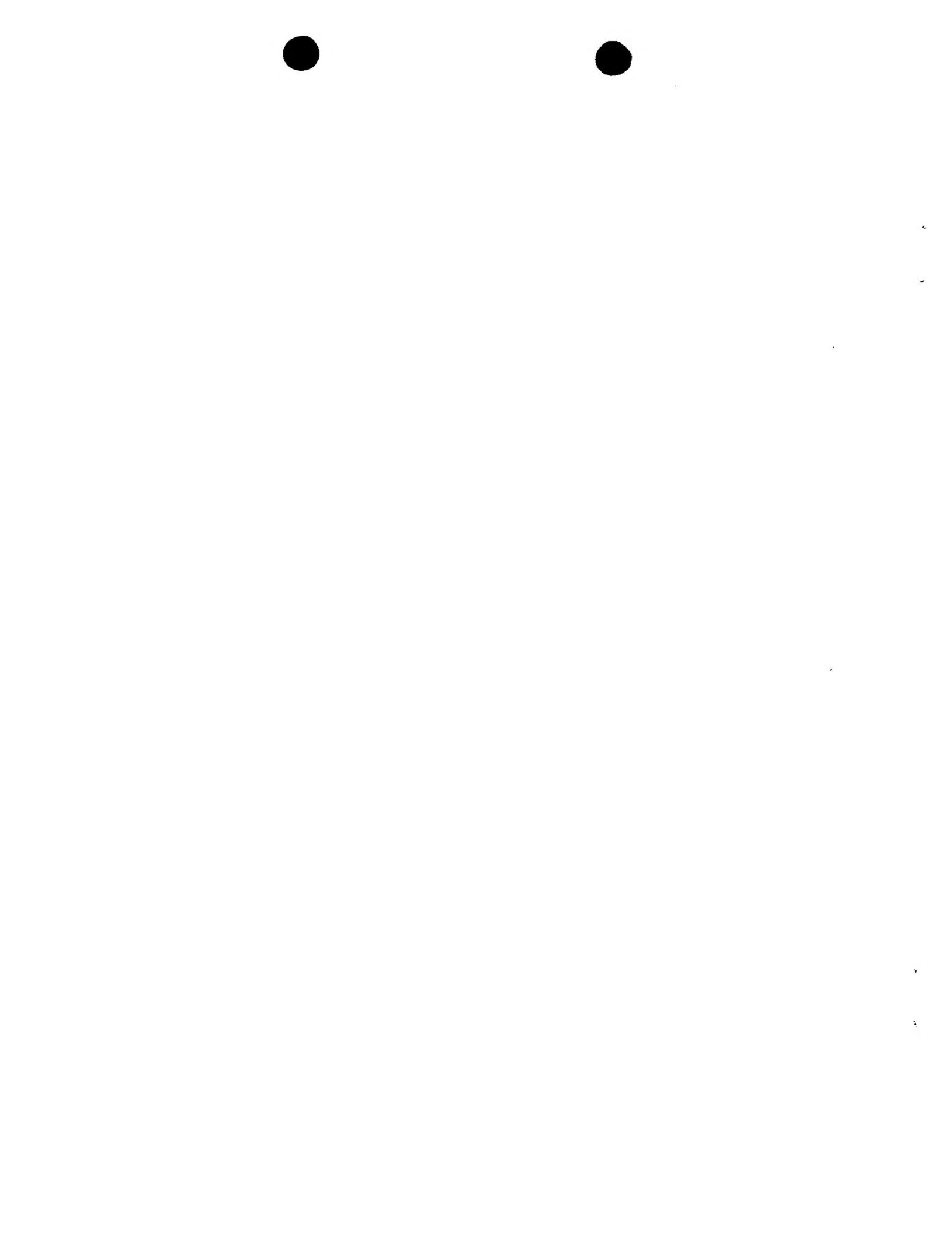


Fig. 3

$$q_i \rightarrow \min \{ \max \{ q_i + \frac{0.25}{\text{threshold}_H - \text{threshold}_N} (Y_i - \text{threshold}_N) - 0.1, 0.1 \}, 1.0 \}$$



INTERNATIONAL SEARCH REPORT

National Application No

PCT/EP 00/00317

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F9/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EVANS D J ET AL: "DYNAMIC LOAD BALANCING USING TASK-TRANSFER PROBABILITIES" PARALLEL COMPUTING, vol. 19, no. 8, 1 August 1993 (1993-08-01), pages 897-916, XP000385007 page 900, paragraph 3.1 -page 902 page 902, paragraph 3.2 -page 903 -----	1,25
A	EP 0 715 257 A (BULL SA) 5 June 1996 (1996-06-05) the whole document -----	1,25

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

26 May 2000

Date of mailing of the international search report

05/06/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Michel, T

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 00/00317

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0715257	A 05-06-1996	FR WO JP US	2727540 A 9617297 A 10507024 T 5993038 A	31-05-1996 06-06-1996 07-07-1998 30-11-1999

INTERNATIONALER RECHERCHENBERICHT

I. nationales Aktenzeichen

PCT/EP 00/00317

A. KLASSEFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G06F9/46

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHEN UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	EVANS D J ET AL: "DYNAMIC LOAD BALANCING USING TASK-TRANSFER PROBABILITIES" PARALLEL COMPUTING, Bd. 19, Nr. 8, 1. August 1993 (1993-08-01), Seiten 897-916, XP000385007 Seite 900, Absatz 3.1 -Seite 902 Seite 902, Absatz 3.2 -Seite 903	1,25
A	EP 0 715 257 A (BULL SA) 5. Juni 1996 (1996-06-05) das ganze Dokument	1,25



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besondere bedeutsam anzusehen ist

"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung,

Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kolliert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erforderlicher Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erforderlicher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann nahelegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

26. Mai 2000

Anmeldedatum des internationalen Recherchenberichts

05/06/2000

Name und Postanschrift der Internationalen Recherchenbehörde

Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Michel, T

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP 00/00317

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
EP 0715257 A	05-06-1996	FR 2727540 A	31-05-1996
		WO 9617297 A	06-06-1996
		JP 10507024 T	07-07-1998
		US 5993038 A	30-11-1999

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)